REVIEW ARTICLE

Deglycosylation of glycoproteins with trifluoromethanesulphonic acid: elucidation of molecular structure and function

Albert S. B. EDGE[1]

Harvard Medical School and Eaton Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, U.S.A.

The alteration of proteins by post-translational modifications, including phosphorylation, sulphation, processing by proteolysis, lipid attachment and glycosylation, gives rise to a broad range of molecules that can have an identical underlying protein core. An understanding of glycosylation of proteins is important in clarifying the nature of the numerous variants observed and in determining the biological roles of these modifications. Deglycosylation with TFMS (trifluoromethanesulphonic acid) [Edge, Faltynek, Hof, Reichert, and Weber, (1981) Anal. Biochem. **118**, 131–137] has been used extensively to remove carbohydrate from glycoproteins, while leaving the protein backbone intact. Glycosylated proteins from animals, plants, fungi and bacteria have been deglycosylated with TFMS, and the most extensively studied types of carbohydrate chains in mammals, the N-linked, O-linked and glycosaminoglycan chains, are all removed by this procedure. The method is based on the finding that linkages between sugars are sensitive to cleavage by TFMS, whereas the peptide bond is stable and is not broken, even with prolonged deglycosylation. The relative susceptibility of individual sugars in glycosidic linkage varies with the substituents at C-2 and the occurrence of amido and acetyl groups, but even the most stable sugars are removed under conditions that are sufficiently mild to prevent scission of peptide bonds. The post-translational modifications of proteins have been shown to be required for diverse biological functions, and selective procedures to remove these modifications play an important role in the elucidation of protein structure and function.

Key words: carbohydrate, deglycosylation, glycoprotein, post-translational modification, protein structure, trifluoromethanesulphonic acid (TFMS).

INTRODUCTION

The elucidation of the multiple products of single genes remains one of the most formidable problems faced in defining the proteome of a tissue. Understanding the contribution of glycans to the diverse structures of glycoproteins is a key aspect of the definition of post-translational modifications, and a way of selectively removing glycans from glycoproteins can be extremely useful in determining the structure of the glycoprotein and its relationship to a known gene. There are now numerous examples in the literature of the removal of carbohydrate groups by TFMS (trifluoromethanesulphonic acid) to produce the intact polypeptide, as shown by molecular mass determination, amino acid sequencing and immunoreactivity with antibodies that are specific for the protein core. Deglycosylation has been used as a method for the identification of glycosylation sites in elucidation of the structure of glycosylated proteins. The procedure has been employed as a probe for the determination of the types of sugar chains added during biosynthesis of a glycoprotein. The TFMS procedure has the advantage over enzymic removal of sugars that all carbohydrate groups tested to date can be removed in a single step, and the procedure can therefore be used in the absence of information about the glycans and their attachment to the protein. This has permitted the use of the TFMS deglycosylation procedure to identify recognition sites for antibodies and bacterial adhesion proteins, and to determine the importance of sugars for biological functions of enzymes, hormones and lectins.

PROCEDURE FOR DEGLYCOSYLATION

Chemical and enzymic procedures can be used in conjunction in deglycosylating glycoproteins, but for many purposes the chemical procedure, which is simple and requires no expensive reagents, is preferable. TFMS-mediated deglycosylation [1] has the advantage that it removes carbohydrate chains from glycoproteins regardless of linkage and composition, and this has been shown in numerous studies using the method.

The deglycosylation procedure removes all sugars from proteins that contain covalently bound sugar chains. The method has been used successfully with O-linked oligosaccharides attached to a serine or threonine residue of a glycoprotein [1–3], N-linked glycans with *N*-acetylglucosamine–asparagine linkage [1,4–7], glycosaminoglycans attached to the core protein of proteoglycans [1,8–10] and collagen saccharides comprising glucose and galactose attached via a hydroxyproline linkage [11–13]. Saccharides have also been removed from glycoconjugates of trematode parasites [14], insect glycoproteins [15,16], sponge proteoglycans [17], galactose–serine linkage and arabinose–hydroxyproline linkage in plants [18–21], mannose linkage to serine in yeast [22], glucose–tyrosine linkage in bacteria [23] and various types of glycoconjugates in archaebacteria [24,25].

Therefore the method can be applied to glycoproteins of unknown structure and used in conjunction with enzymes or physical methods to characterize the glycosylation and the underlying protein. Amide bonds are stable to TFMS, and the difference

Abbreviations used: a.m.u., atomic mass units; CEA, carcinoembryonic antigen; endo F and endo H, endo-β-*N*-acetylglucosaminidase F and H respectively; MALDI–TOF, matrix-assisted laser-desorption ionization–time-of-flight; TFMS, trifluoromethanesulphonic acid.

[1] e-mail Albert_Edge@meei.harvard.edu

in stability of the glycosidic bond as compared with the peptide bond is the basis of the procedure, as proteins have repeatedly been demonstrated to be stable to the reagent under the conditions used to remove sugars. The exception is the innermost asparagine-linked *N*-acetylglucosamine of N-linked chains, which is attached to the protein by an amide bond and is stable to TFMS. The reaction is carried out under anhydrous conditions in TFMS without a solvent in the presence of a scavenger that protects the amino acids, and proceeds by solvolytic cleavage of glycosidic bonds, leaving the protein intact. The solvolysis should be performed in the absence of water to prevent the introduction of hydrolytic conditions which could result in cleavage of peptide bonds. TFMS is used as a reagent for cleavage of the peptide from the resin and removal of protecting groups in the solid-phase synthesis of peptides [26,27]. The peptide bond is stable and the side chains of amino acids that could undergo reactions with intermediates of the deblocking steps are protected by the use of scavengers in the reaction. Thus in the presence of anisole any reactions involving tyrosine, tryptophan, cysteine, methionine, serine and threonine are limited to a low level. The same methods are used for deprotection with HF (hydrogen fluoride), although TFMS is a more powerful deblocking agent and is less toxic [28,29].

Apart from *N*-acetylglucosamine attached to an asparagine residue, conditions can be found to remove most sugars from glycoproteins. The innermost *N*-acetylglucosamine of N-linked sugar chains was seen to be stable to solvolysis when glycoproteins such as fetuin were treated with the reagent at both 25 °C and 0 °C for prolonged periods of time [1]. Unlike this sugar, which has an amide linkage to an asparagine residue, all sugars in glycosidic linkage were sensitive to the reagent and could be completely removed. The susceptibility of glycosidic bonds to TFMS varies with the nature of the sugar and its substituents. The innermost sugar of O-linked saccharide chains, *N*-acetylgalactosamine linked to a serine or threonine residue, is one of the most resistant linkages found in mammalian glycoproteins, and this has also been observed with HF [30]. Whereas all galactose and mannose chains are removed from fetuin at 0 °C or 25 °C (removal of these sugars is complete in less than 1 h at 0 °C), *N*-acetylgalactosamine is fairly stable at 0 °C in TFMS and its rate of loss is low at 25 °C (75 % at 3 h). Therefore, deglycosylation under conditions that do not remove the serine/threonine-linked *N*-acetylgalactosamine completely will yield an extensively deglycosylated mucin if most of the sugars are external to the *N*-acetylgalactosamine, as found in MUC1 [31–33], whereas in a mucin with short sugar chains, such as submaxillary mucin with predominantly disaccharides [34], the percentage of sugars removed will be lower (50 %, if all *N*-acetylgalactosamine–serine/threonine remains). Very heavily glycosylated proteins, including proteoglycans [9,10,35–37] and plant glycoproteins [17–19,21], have been completely deglycosylated by TFMS. The complete deglycosylation of glycoproteins that contain as much as 80 % carbohydrate indicates that any difficulty encountered with mucins is based on structural features, not the extent of glycosylation. The use of conditions that retain *N*-acetylgalactosamine attached to serine and threonine residues has allowed the determination of linkage sites in mucins. Removal of all sugars, apart from the internal *N*-acetylgalactosamine, has been used to identify the serine and threonine residues to which sugars are attached.

The scavenger most frequently used for TFMS treatment has been anisole, although other reagents have also been included. In the presence of anisole as a scavenger the amino acid side chains are recovered intact. By providing a sink for reactive electrophiles the scavenger protects the side groups, whereas in the absence of scavenger, when peptides are deprotected with HF or TFMS, reactions involving serine, threonine, aspartate, cysteine, methionine and the aromatic amino acids have been described [27,28]. Other investigators have suggested that results are improved with the use of phenol or toluene instead of anisole [38,39]. An effect on the rate of deglycosylation is apparent with increased anisole concentrations [31,39,40], and the highest rates are possible in the absence of scavenger. The 1:2 ratio of anisole to TFMS has been reduced without compromising stability of side chains, and a concentration of 10 % anisole or phenol is adequate. In combination with reduction of the size of the sample, the volume of the reaction can be scaled down to less than 100 μl. The temperature is kept at 0 °C unless O-linked *N*-acetylgalactosamine must be completely removed, in which case the reaction is carried out at a higher temperature.

COMPARISON WITH ENZYMIC DEGLYCOSYLATION

As compared with enzymic removal of carbohydrate chains from glycoproteins, the TFMS deglycosylation procedure has the advantage that all glycans can be removed from a glycoprotein regardless of structure, and this provides ready access to the protein core. In studies of well-characterized glycoproteins, chemical deglycosylation can be the preferred method if carbohydrate chain removal is difficult to achieve with the use of enzymes. This can be encountered if the specificity of an enzyme does not allow cleavage of a certain saccharide due to a bad fit of the sugars into an active site or if an appropriate enzyme for a particular type of linkage is not available. Enzymes available for N-linked oligosaccharides of mammalian glycoproteins include N-glycanase [peptide-*N*-(*N*-acetyl-β-glucosaminyl)asparagine amidase] and endo F and endo H (endo-β-*N*-acetylglucosaminidases F and H). If the types of glycosylation of a glycoprotein are unknown, their susceptibility to these enzymes is uncertain and resistant asparagine-linked saccharides are sometimes encountered, particularly in plants [39]. The enzymes for removal of N-linked chains usually cleave the commonly found types of N-linked oligosaccharides in mammals, but may require detergent or a denaturing agent for access of the enzyme to the bond to be cleaved to completely deglycosylate a protein [41,42]. This can be a disadvantage if the biological activity of the protein is to be tested after deglycosylation.

Enzymes available for O-linked chains are extremely limited. The commercially available enzyme O-glycanase has a specificity restricted to the disaccharide sequence Gal-GalNAc only [42], and, therefore, resistance to O-glycanase should not be taken as evidence for the lack of O-linked saccharide chains. Thus O-glycanase has not been used successfully to remove oligosaccharide chains from most mucins due to the complexity of the structure of the O-linked oligosaccharides. Except for the internal *N*-acetylgalactosamine, whose glycosidic linkage is more resistant to TFMS than other bonds between sugars in glycoproteins, removal of all other sugars of O-linked glycan is efficiently achieved with TFMS in both mucins and glycoproteins with O-linked oligosaccharides [1] and therefore the extent of deglycosylation is nearly complete [34,40,43,44]. However, for some studies removal of this sugar is not required and for the investigation of the glycosylation of individual amino acids, deglycosylation is performed for short periods at 0 °C, because retention of a single *N*-acetylgalactosamine is desired [31,45].

Although the protein core is not obtained, removal of glycosaminoglycan chains from proteoglycans has been performed with heparanases, chondoitinases, hyaluronidase and keratanases [46]. These enzymes have been useful for the characterization of the

repeating disaccharide units that are released by digestion, but information about the types of chains that are present must be available before the correct enzyme is chosen for removal of carbohydrates from the molecule. Carbohydrate removal by the glycosaminoglycan-degrading enzymes is always incomplete because the core saccharide GlcA-Gal-Gal-Xyl, which serves as an attachment site for the glycosaminoglycan chains, is not removed by these enzymes. In addition, proteoglycans with a homogeneous glycosaminoglycan chain are an exception, as most proteoglycan core proteins have several types of carbohydrate chain attached which can only be removed with multiple enzymes [10,36,37]. The chemical deglycosylation procedure removes all types of glycosaminoglycan chains [1], and this includes the linkage tetrasaccharide that is left by the glycosaminoglycan-degrading enzymes [37]. The glycosylation of collagens and proteins with collagen-like domains at hydroxyproline is not susceptible to cleavage by a known specific endoglycosidase. The carbohydrate attached to collagen at hydroxylysine can be removed with TFMS [11,12]. No reports of TFMS treatment of glycoproteins with *O*-fucosyl linkage of sugar chains to serine or threonine residues [47] have been published to date.

Again, in studies of proteins of unknown structure the ability to remove all modifications that alter the electrophoretic behaviour or mass spectral fragmentation of a protein is desirable. Thus the chemical procedure has a number of advantages over enzymic deglycosylation.

DETERMINATION OF MOLECULAR MASS

Removal of the carbohydrate groups from a glycoprotein permits the molecular mass of the protein portion to be determined, and deglycosylation with TFMS has been widely applied for this purpose. This can be useful both for the elucidation of the post-translational modifications and for the correlation of the molecular mass to the predicted size from the cDNA sequence. Numerous examples of discrepancies between the predicted molecular mass from the cDNA and the observed molecular mass of the protein, taking into account the removal of the leader peptide, have been resolved based on the conversion to the predicted molecular mass after TFMS treatment [9,17,19,48–51]. The size difference can be even larger than expected and the effect of chemical deglycosylation more striking than expected due to the heterogeneity of carbohydrate chains and the anomalous behaviour of glycoproteins on SDS gels. Thus proteins with multiple forms will be reduced to a single protein if there are no proteolytically processed forms of the protein.

In several instances in which MS has been used for the determination of molecular mass after deglycosylation, the masses that have been found have corresponded exactly with the mass expected from the amino acid sequence. A fungal glycoprotein, hydrophobin SC3, was deglycosylated with TFMS to yield a protein with a mass based on MALDI–TOF (matrix-assisted laser-desorption ionization–time-of-flight) MS of 10 854 Da, as compared with 14 400 Da for the fully glycosylated protein and 10 856 Da for the theoretical average mass [52]. *Mycobacterium tuberculosis* antigen was reduced to a molecular mass of 28 782 Da after deglycosylation [50], corresponding to the molecular mass deduced from the cDNA sequence. Glycoforms with between 1 and 9 mannose residues were apparent in the MS profiles of the intact protein, but were absent after TFMS deglycosylation. *Toxoplasma gondii* surface antigen I expressed in *Pichia pastoris* had a series of glycoforms of molecular masses of 29, 38/45, and 50/60 kDa that were all reduced to 29 kDa after TFMS treatment [22]. MALDI–TOF and ESI (electrospray ionization) MS analysis of the 29 kDa protein revealed species of 28 694, 28 856, 29 017, 29 180 and 29 342 a.m.u. (atomic mass units), which were all reduced to 28 694 a.m.u. after chemical deglycosylation.

SEQUENCING AND IDENTIFICATION OF SITES OF GLYCOSYLATION

The sequence of a glycosylated protein can be difficult to determine because the protein is less susceptible to proteases that are used to produce peptides for sequencing, and the glycosylated amino acid is lost during processing of the phenylthiohydantoin obtained in the Edman degradation. In sequencing by MS the heterogeneity in the size of the glycosylated amino acid complicates the ions obtained during fragmentation. The removal of carbohydrate can help with both of these problems. Studies that have determined the sequence of a TFMS-treated protein indicate that the deglycosylation does not alter recovery of the individual amino acids. Comparison of the sequence with and without deglycosylation can provide the identity of the amino acids that are lost during sequencing of the native protein, and the linkage sites of oligosaccharides attached at asparagine residues are obtained as a single *N*-acetylglucosamine attached to an asparagine residue. This *N*-acetylglucosamine–asparagine derivative can be separated from other phenylthiohydantoin amino acids by HPLC and could be used to identify the sites of asparagine-linked chains in carcinoembryonic antigen (CEA) [6,53]. Thus, by using the innermost *N*-acetylglucosamine as a reporter group, TFMS treatment can be useful in determining the sites of glycosylation on a glycoprotein. Chymotryptic peptide maps of the protein after deglycosylation by TFMS corresponded with predicted peptides from the cDNA [6], and fast-atom-bombardment MS indicated that the peptides which contained glycosylation sites were larger than expected from the sequence by a single *N*-acetylglucosamine. A mass ion 203.1 a.m.u. higher than predicted for the peptide was obtained, and thereby confirmed the sites of glycosylation.

Investigations of mucins deglycosylated with TFMS have allowed the isolation of tandem repeats that are the sites of O-glycosylation which were not digested by proteases before deglycosylation. An 81-residue tandem repeat could be obtained from porcine submaxillary mucin by digestion with trypsin after TFMS treatment to remove sugars [45], and a 20-residue tandem repeat was produced by clostripain digestion of deglycosylated MUC1 [31]. This repeat structure could not be obtained by proteolytic digestion of the intact mucin molecule which was protease resistant. By using mild conditions of TFMS treatment (1 h at 4 °C) for porcine submaxillary mucin [45], it was possible to cleave all sugar linkages, while leaving the core *N*-acetylgalactosamine attached, and this *N*-acetylgalactosamine–serine or *N*-acetylgalactosamine–threonine linkage could be detected in Edman sequencing, allowing identification of glycosylation sites. A similar approach was taken with MUC1, except that TFMS treatment was carried out for 20 min without anisole. MALDI–TOF MS allowed the identification of glycosylation sites by the increment in mass due to the monosaccharide of 203.2 a.m.u. attributed to the peptides with single *N*-acetylgalactosamine residues attached at a serine or threonine residue [31]. This was confirmed by Edman sequencing which demonstrated the *N*-acetylgalactosamine–serine and *N*-acetylgalactosamine–threonine residues at sites of O-linked oligosaccharide attachment (Figure 1). In a method for complete deglycosylation of mucins the serine- or threonine-linked *N*-acetylgalactosamine was removed by periodate oxidation after TFMS treatment [34].

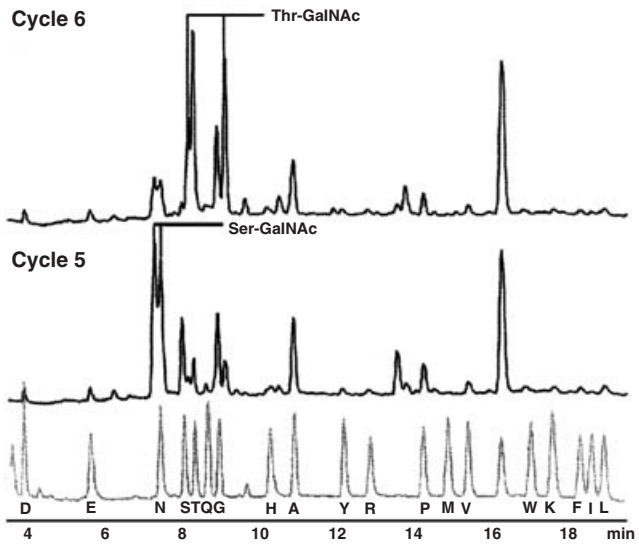

Figure 1 Identification of serine/threonine-linked glycosylation sites

Edman degradation of MUC1 [31] after TFMS deglycosylation under conditions selected to retain the internal *N*-acetylglucosamine of O-linked glycans. The results of cycles 5 and 6 are shown. A GalNAc–serine linkage was identified at cycle 5 and a GalNAc–threonine linkage at cycle 6 corresponding to O-linked glycans at positions 5 and 6 of MUC1. Reproduced, with permission, from Muller, S., Goletz, S., Packer, N., Gooley, A., Lawson, A. M. and Hanisch, F. G. (1997), J. Biol. Chem. **272**, 24780–24793, © The American Society for Biochemistry & Molecular Biology.

After deglycosylation of *Geobacillus stearothermophilus* S-layer glycoproteins, a single protein band of 93 kDa was produced from the initial four proteins of 93, 119, 147 and 170 kDa [24]. This band corresponded well to the calculated molecular mass of 93 684 Da from the amino acid sequence. Edman sequencing of peptides produced by papain treatment and then subjecting them to deglycosylation revealed that this protein had sites of sugar attachment at Thr^{620} and Ser^{794}. The TFMS deglycosylated peptides had serine and threonine residues at these positions, while the native peptides had blanks at the positions of glycosylation. The linkage in these glycoproteins was through galactose in a β configuration to the serine and threonine residues. This indicated for the first time the nature of the linkage of these bacterial glycoproteins, and demonstrates how chemical deglycosylation can be used to establish linkage sites in a glycoprotein.

Treatment of inter-α-trypsin inhibitor with TFMS allowed the sequence of the protein to be determined, and indicated that serine was present at residue 10 which had previously yielded no amino acid identity [54]. This amino acid was not found after hyaluronidase treatment [54], because the serine is still glycosylated after enzymic treatment due to the incomplete removal of the glycosaminoglycans by enzymic treatment.

DEGLYCOSYLATION TO DISTINGUISH GLYCOSYLATION VARIANTS FROM PROTEIN DIFFERENCES

The complete removal of carbohydrate from proteins has been a powerful tool in the analysis of post-translational modifications of proteins, and particularly in determining whether changes observed in size are due to proteolytic processing, alternative splicing, addition of several types of carbohydrate chains or elongation of existing carbohydrate chains in secreted and cell surface proteins. A greater decrease in size after TFMS treatment as compared with N-glycanase-catalysed cleavage can indicate the occurrence of O-linked sugars. In this subtractive analysis TFMS removes both N-linked and O-linked sugar chains, whereas O-linked sugars are resistant to N-glycanase.

Several investigators have studied zona pellucida glycoproteins by using TFMS treatment to remove the carbohydrate chains [2,51,55]. The zona pellucida glycoprotein ZP3 was shown to be the egg receptor for sperm [2], and the activity was specifically assigned to the O-linked glycans by the loss of sperm-binding activity after TFMS treatment [2]. The molecular masses of TFMS-treated ZP2 and ZP3 were 67 kDa and 37 kDa respectively after removal of the O-linked and N-linked oligosaccharides [51]. This was smaller than the sizes of 75 kDa and 44 kDa based on predicted sequences from the cDNA. This revealed that the processed forms of ZP2 and ZP3, which are made in the growing oocyte and secreted before forming a protective coat around the egg, are proteolytically processed, as well as being glycosylated, at asparagine and at serine/threonine residues. Thus with proteolysis and glycosylation required for production of the final protein, the post-translational modifications were complex, and TFMS deglycosylation was able to distinguish between the two types of modifications. Consensus sites for cleavage by furin were used in this protein during the transit of the proteins through the Golgi, and this was confirmed by C-terminal analysis of the final protein. Other studies with ovarian glycoproteins have shown that the α and β forms of oviductin contain the same protein core, as revealed by TFMS treatment, which had no effect on the α form and converted the β form to the same mass and pI as the α form [56].

Deglycosylation with TFMS has been used for elucidation of the structure of cell surface receptors. In a study of the insulin receptor, the difference between the molecular mass after N-glycanase treatment and the mass after TFMS treatment suggested the occurrence of O-linked oligosaccharides on the β subunit [3]. Endo-β-galactosidase treatment indicated that the N-linked chain contained repeating poly(*N*-acetyl-lactosamine) units [57]. TFMS treatment reduced the apparent molecular mass of the angiotensin II receptor from 65 kDa to 40 kDa, which agreed with the calculated mass of the cloned receptor of 40 889 Da [48].

The core proteins of proteoglycans had been difficult to identify prior to the use of chemical deglycosylation [8]. The major challenge has been to clarify the number and molecular mass of protein cores. A single proteoglycan gene can give rise to molecules that differ in the size and number of sugar chains attached, the size of the core protein and the mode of attachment to the membrane, which can be through a transmembrane sequence or a glycosylphosphatidylinositol anchor. Comparison to enzymic deglycosylation was used to demonstrate that the molecular mass of the core protein of glomerular basement membrane proteoglycan was 128 kDa [8], and that the intact molecule, in addition to the heparan sulphate chains, had N-linked and O-linked oligosaccharides (Figure 2). This became apparent from the difference between the partially deglycosylated protein obtained by a combination of heparitinase and N-glycanase treatments and the completely deglycosylated protein obtained after TFMS treatment. van den Heuvel et al. [58] found two core proteins of 160 kDa and 110 kDa after TFMS deglycosylation of human tubular basement membrane proteoglycans of molecular mass 200–400 kDa, whereas bovine tubular basement membrane was found to have smaller core proteins after deglycosylation [35]. The proteoglycans of glomerular and tubular basement membranes are now thought to correspond to agrin [59,60] and perlecan [59,61], proteoglycans that are found in basement membranes of a number of tissues, but, if so, some of the core proteins that have been identified must be proteolytically processed. TFMS treatment of the 150–300 kDa heparan sulphate

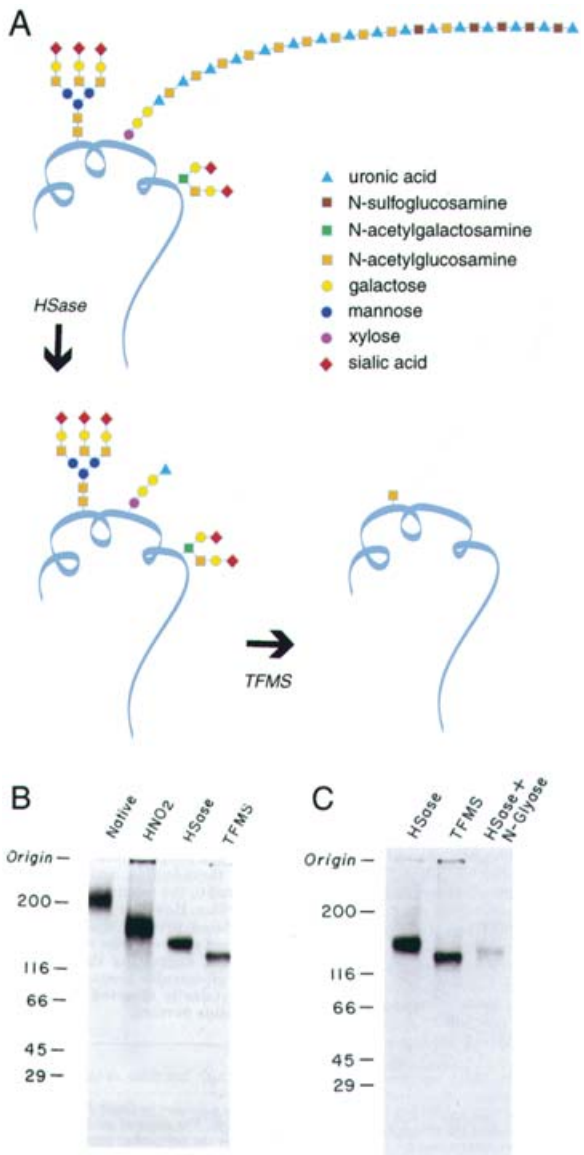

Figure 2 Comparison of enzymic and TFMS deglycosylation of a proteoglycan-containing N-linked and O-linked glycans and heparan sulphate chains

(**A**) The scheme indicates all three types of carbohydrate chain attached to a core protein (blue-grey line). For the illustration, the heparan sulphate chain is depicted shorter than the actual glycosaminoglycan found on the glomerular basement membrane proteoglycan, which is approx. 60 sugars in length. The uronic acids are iduronic and glucuronic acids. The outer third of the polymer contains extensive sulphation, whereas the inner two thirds consists predominantly of GlcA–GlcNAc disaccharides [121]. The proteoglycan can be deglycosylated incompletely with heparitinase (HSase), yielding a proteoglycan with the linkage tetrasaccharide and N-linked and O-linked chains, or it can be treated with TFMS to deplete it of all sugars, except the single *N*-acetylglucosamine of the N-linked glycan bound to asparagine [8]. (**B**) Deglycosylation of the heparan sulphate proteoglycan of the glomerular basement membrane with TFMS reveals that the core protein is 128 kDa. An intermediate size is observed for the heparitinase-treated proteoglycan and for the nitrous acid-treated molecule, which is truncated by cleavage at *N*-sulphoglucosamines [brown squares in (**A**)] in the distal portion of the glycosaminoglycan chains. (**C**) Partial deglycosylation by N-glycanase (N-Glyase) and heparitinase leaves the linkage portion of the glycosaminoglycan chains and the O-linked glycans intact, which accounts for the difference between the enzymically treated molecule and the core protein (TFMS). Reproduced, with permission, from Edge, A. S. B. and Spiro, R. G. (1987), J. Biol. Chem. **262**, 6893–6898, © The American Society for Biochemistry & Molecular Biology.

proteoglycan isolated from the extracellular matrix of rat liver revealed a core protein of 40 kDa [62]. This proteoglycan was present in the basement membrane of the perisinusoidal space of Disse, closely associated with the sinusoidal cells. This deglycosylation result indicates that the proteoglycan is distinct from the proteoglycan of the glomerular basement membrane, although the molecular mass of the intact molecule is similar.

Chondroitin sulphate proteoglycans secreted from B-cells had molecular masses of 90 kDa, 130 kDa and 150 kDa from three cell lines, but all could be reduced to a molecular mass of 21 kDa after TFMS treatment [37]. The chemical deglycosylation removed chondroitin and heparan sulphate chains, and this included the GlcA-Gal-Gal-Xyl linkage saccharide which was not removed by treatment with heparitinase. A corneal proteoglycan with keratan sulphate chains was shown to have a core protein of 35 kDa [9]. Lewis lung carcinoma expresses a proteoglycan with both heparan and chondroitin sulphate chains and a core protein of 40 kDa [36]. Binding of the lung cells to the matrix is mediated by the affinity of the heparan sulphate chains for fibronectin, and this affinity was shown to be localized in the heparan sulphate chains by a combination of chemical and enzymic deglycosylation.

ELUCIDATION OF PROCESSING AND TRANSPORT THROUGH INTRACELLULAR ORGANELLES BY THE USE OF DEGLYCOSYLATION

The combination of chemical and enzymic deglycosylation is useful for defining the types of substituents attached to a protein. Multiple types of carbohydrate chain are added to many glycoproteins, and a difference in the results of deglycosylation by TFMS, as compared with enzymes, can yield insights into the structure of the attached sugars. Thus, for example, the resistance of a glycoprotein to an equal extent of deglycosylation by N-glycanase and TFMS indicates that O-linked carbohydrate chains are present, and the acquisition of this resistance is observed as the protein is processed in subcellular organelles.

Deglycosylation with TFMS revealed that intestinal lactase-phlorizin hydrolase was present in the intestine in several glycoforms that differed by the addition of O-linked glycans to the glycoprotein [63]. Treatment with N-glycanase yielded a 118 kDa and a 126 kDa protein from the 160 kDa mature protein, whereas TFMS treatment yielded only the 118 kDa protein, indicating that the 126 kDa protein contained O-linked sugars. This was confirmed by binding of the 126 kDa form to *Helix pomatia* lectin. The demonstration of the presence of O-linked sugars on the 126 kDa form of lactase-phlorizin hydrolase showed that the difference in mass of the two bands was due to glycosylation variants and not to a difference in the protein resulting from proteolytic processing or alternative splicing. The presence of the O-linked chains increased the activity of the enzyme by raising the V_{max} without altering the K_m for lactose, and the addition of both N- and O-glycosidically linked chains was thought to be governed by the position of the mucosal cell in the intestinal villus, implying that these modifications played a role in the regulation of the enzyme's activity.

Further studies of glycoproteins secreted into the intestine showed that an immature form of dipeptidyl peptidase IV (molecular mass = 100 kDa) was susceptible to endo H and was reduced to the same mass species (molecular mass = 85 kDa) by either endo H or TFMS (Figure 3) [64], whereas the mature form of the enzyme with complex chains (molecular mass = 124 kDa) was resistant to endo H. This form of the enzyme was converted to the 85 kDa core protein by TFMS, but only to a molecular mass of 90–110 kDa by N-glycanase

treatment [64], indicating that the mature protein had been further glycosylated with O-linked chains. The N-linked chains are attached to proteins in the endoplasmic reticulum and the results showed that both the processing of the N-linked chains and the addition of the O-linked oligosaccharides of dipeptidyl peptidase occurred in the Golgi. The acquisition of resistance to complete deglycosylation by N-glycanase occurred simultaneously with the loss of susceptibility to endo H, confirming that the processing of N-linked oligosaccharide chains to complex forms, which accounts for resistance to endo H, was temporally linked to the addition of O-linked sugars (Figure 3), which accounts for the resistance of the glycoprotein to complete deglycosylation by N-glycanase. Thus the comparison of the results of TFMS deglycosylation to those of N-glycanase treatment allows an assignment of the timing of addition of O-linked chains.

DEGLYCOSYLATION OF BACTERIAL AND PLANT GLYCOCONJUGATES

Owing to the lack of enzymes for characterization and removal of plant oligosaccharides and the complexity of linkage types and diversity of carbohydrates synthesized in plants, there has been a need for procedures to aid in the characterization of these glycoconjugates. TFMS deglycosylation has been particularly useful in the characterization of plant and bacterial glycoproteins.

In the trafficking of plant glycoproteins to the cell surface and to intracellular compartments, the use of TFMS for the removal of all carbohydrate chains has clarified the timing of proteolytic processing, as compared with the attachment of sugars in the secretory pathway [65–69]. Both arabinose–hydroxyproline- and galactose–serine-linked sugars were completely removed by TFMS in 3 h at 0 °C from potato (*Solanum tuberosum*) lectin, which is 50 % carbohydrate [18]. These types of linkages are common in plant glycoproteins, and chemical deglycosylation has been applied successfully to remove arabinogalactan- and galactose-rich carbohydrate chains in numerous laboratories [19–21]. TFMS has also been useful for the removal of N-linked chains of plants and bacteria that are resistant to N-glycanase [16,70].

The power of using this procedure to cleave glycosidic linkages without knowledge of structure has proved advantageous in studies of bacterial and fungal glycoproteins. TFMS treatment of glycoproteins from the S-layer of the ancient Archaebacteria and Gram-negative eubacteria has aided in the elucidation of the structure of their surface coats. The finding that carbohydrates were linked to proteins in bacteria was unexpected, as bacteria had not been thought to possess the enzymes for elaboration of glycoproteins, but a large variety of such molecules have now been characterized [24,25] and the occurrence of glycoproteins in bacteria has been confirmed. TFMS treatment of S-layer glycoproteins from *Clostridium thermohydrosulfuricum* resulted in the identification of a core protein of 87.5 kDa and indicated that this was a glycosylated protein [23]. The sugars were attached through a glucose–tyrosine linkage. Proteins from the S-layer of *C. thermosaccharolyticum* [71] were in a heterogeneous range from 83 kDa to 210 kDa and were all reduced to a single band at 83 kDa. Bacterial glycoproteins from *M. tuberculosis* were deglycosylated to yield core proteins of 71 kDa, 60 kDa, and 45 kDa, which could stimulate proliferation of peripheral blood mononuclear cells from infected patients [72] and could be used for the generation of a vaccine that was protective in mice [73]. The deglycosylated 45/47 kDa antigen had decreased immunogenicity when tested for its ability to elicit an immune response in guinea pigs [50]. Treatment of cell wall extract with TFMS released a protein of 23 kDa that was not bound

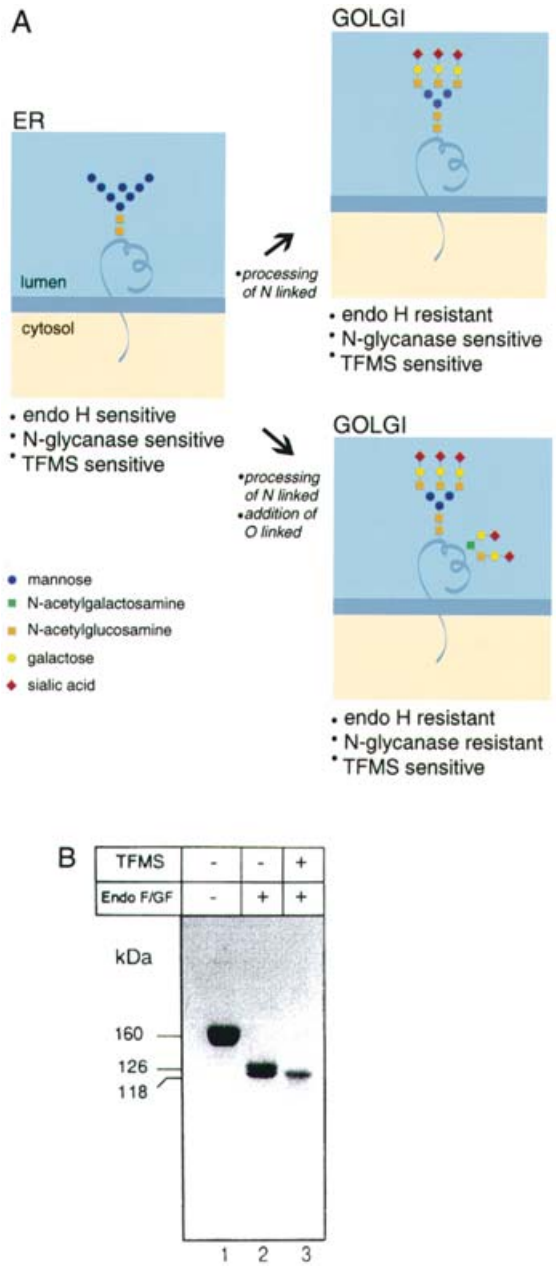

Figure 3 Elucidation of steps in the elaboration of a glycoprotein by partial deglycosylation of biosynthetic intermediates in the pathway

Treatment of the biosynthetic intermediates with TFMS, N-glycanase or endo H deglycosylates the protein partially or fully, yielding products that are characteristic of the stage of glycosylation. In (**A**) the transfer of an N-glycan to the protein in the endoplasmic reticulum (ER) yields a precursor with carbohydrate chains sensitive to endo H, N-glycanase and TFMS. A $Man_8GlcNAc_2$ structure is shown. Upon transport of the glycoprotein to the Golgi the N-glycan is converted to a complex form (upper right-hand panel) that is now resistant to endo H. Proteins that have O-linked chains added (lower right-hand panel) at the time of N-linked saccharide processing are no longer deglycosylated to the same extent by N-glycanase, and this is revealed by comparison with the TFMS treatment which, unlike endo H and N-glycanase, yields the fully deglycosylated product. In (**B**) the products of N-glycanase and endo F treatment of lactase-phlorizin hydrolase are compared with the product of TFMS treatment [63]. The band of 126 kDa was found to contain O-linked glycans that could not be removed with the enzymes, whereas TFMS deglycosylation yielded a 118 kDa product corresponding to removal of both N- and O-linked glycans. Reproduced, with permission, from Naim, H. Y. and Lentze, M. J. (1992), J. Biol. Chem. **267**, 25494–25504, © The American Society for Biochemistry & Molecular Biology.

to peptidoglycan [74]. Gram-positive bacterial cell wall polymers from alkaliphilic *Bacilli* have been deglycosylated with TFMS to yield peptides that are involved in the structure of the wall [75]. A fungal glycoprotein, hydrophobin, has been deglycosylated with TFMS to yield a protein core of 10854 Da [52]. The deglycosylated protein retained the ability of the parent molecule to form an amphipathic interface due to its high biosurface activity, but the deglycosylated protein showed lower tensile strength, indicating a role for the sugars in the formation of the membranes that contribute to the structure of fungal hyphae [52].

Glycosidic bond cleavage by TFMS has been shown to be useful for the structural characterization of bacterial polysaccharides, because it yields characteristic fragments [76]. The glycosidic linkages are cleaved under anhydrous conditions, unlike amide bonds which are stable. Different glycosidic linkages vary in their susceptibility to solvolysis by TFMS, because of the effect of substituents at C-2 and of the occurrence of amido and acetyl groups. Thus the use of TFMS as a reagent can be extended to analysis of the polysaccharides themselves. Rather than simply removing all the carbohydrate chains from a protein to obtain the core peptide, the procedure can be used to obtain fragments or individual sugars (Figure 4). By permitting the reaction to go to completion, some of the unusual monosaccharides of bacterial capsule polysaccharides could be obtained after the cleavage of all glycosidic linkages, allowing structural analysis of the components [77]. Complete fragmentation was more efficient with TFMS than with HF, which has also been used for deglycosylation, but was not as effective for the cleavage of these sugars, and under the correct conditions the more sensitive bonds could be cleaved, whereas the less sensitive bonds were retained [76]. The O-specific polysaccharide of *Vibrio cholerae* 09 was resistant to cleavage by HF, but upon TFMS treatment yielded characteristic fragments in good yields at the diamino sugar residues [78]. TFMS is preferable to HF because of the physicochemical properties of HF (boiling point = 19.6 °C) and because HF is corrosive and hazardous. The order of the stability of glycosidic linkages was 5,7-diamino-3,5,7,9-tetradeoxynonulosonic acids > 2,3-diamino-2,3-dideoxyuronic acids >uronic acids∼2-amino-2-dideoxyhexoses∼2,4-diamino-2,4,6-trideoxyhexoses > 2-amino-2,6-dideoxyhexoses > hexoses, and this was useful for the structural elucidation of these polymers.

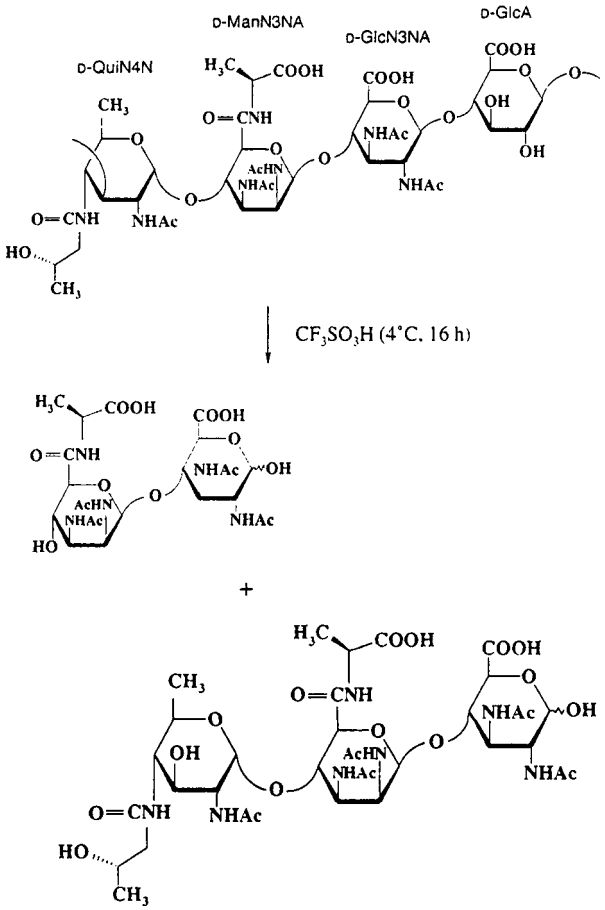

Figure 4 Partial cleavage of glycosidic bonds can be used to obtain characteristic fragments of polysaccharides that provide structural information

A selective cleavage of polysaccharides from bacteria was observed using TFMS under the conditions indicated [76]. The polysaccharide from *Pseudomonas* sp. KMM 634 was not cleaved at $ManN_3NA$, but was cleaved completely at the glycosidic linkages of GlcA and $GlcN_3NA$ and partially at the linkage of $QuiN_4N$ to yield the di- and tri-saccharides shown. The selective cleavage is used to produce these fragments from which the structure of the polysaccharide is obtained. Reproduced with permission of CSIRO Publishing. Originally published as Figure 2 in the Australian Journal of Chemistry, vol. 55 (1–2): 69 (Knirel, Y. A. and Perepelov, A. V.), © CSIRO 2002.

CAN TFMS TREATMENT BE USED TO IDENTIFY THE MODIFICATION OF A PROTEIN?

An important question for the use of this procedure as a tool for characterization of newly discovered glycoproteins is whether the effect of TFMS on a glycoprotein is sufficiently specific that a change in molecular mass after treatment can be ascribed to removal of oligosaccharides. That is, if the molecular mass is decreased by TFMS treatment, can the difference in size be attributed to removed sugars or are other post-translational modifications altered by the chemical treatment? The accumulated evidence suggests that post-translational modifications other than glycosylation are stable to TFMS treatment.

This has been tested for the modification of proteins by both sulphate and phosphate. In a study of myelin-associated glycoprotein labelled with ^{32}P, the label was retained after TFMS treatment [79] and was present in phosphoserine, phosphothreonine and phosphotyrosine residues. This indicates that phosphate on amino acids is stable to the treatment and that the procedure can be used to distinguish between phosphorylation and glycosylation. Sulphated saccharides on α-amidating mono-oxygenase-1 were removed by TFMS, whereas sulphated tyrosine in the C-terminal region of α-amidating mono-oxygenase-4 was stable to treatment [80]. Sulphated prolactin was also stable to treatment, indicating that sulphate was present on tyrosine residues [81]. Thus deglycosylation with TFMS can be used to distinguish between sulphation or phosphorylation of the polypeptide (retained after treatment) and sulphation or phosphorylation of sugars (lost after treatment). The retention of phosphate and sulphate is also useful because the ^{32}P and ^{35}S labels can be used to follow a protein before and after TFMS deglycosylation [81].

Haem groups are stable to TFMS treatment, and Ca^{2+} also remains stably attached after deglycosylation. In a study of horseradish peroxidase, the choice of protocol for purification of the protein was shown to be important for retention of the haem groups, which were extracted by aqueous pyridine used for neutralization of the acid during the purification steps, but could be retained by using Tris base for neutralization [39]. Glycosylphosphatidylinositol anchors are removed, as expected,

by cleavage at the glycosidic linkages in the oligosaccharide chain [82–84]. Fatty-acid acyl groups are stable to the treatment and can therefore be discriminated from glycosylation as an explanation for a change in size after translation. Cystine disulphides are unaffected by TFMS, because multimeric proteins remain intact after treatment [10,44,85].

Glycoforms of many glycoproteins have been resolved and deconvoluted to a single protein by the conversion to a deglycosylated core protein after TFMS treatment [9,17,19,48–51]. The stability of the individual amino acids and of post-translational modifications makes the procedure amenable to the identification of glycosylation on a previously uncharacterized glycoprotein.

DETERMINATION OF THE REACTIVITY OF ANTIBODIES WITH PROTEIN OR CARBOHYDRATE EPITOPES AFTER DEGLYCOSYLATION

Deglycosylation of glycoproteins has been useful in solving the question of the relative contribution of carbohydrates to antigenic epitopes in glycoproteins. Antibodies against glycoproteins can react with carbohydrate epitopes, peptide epitopes or combined epitopes that reside in the protein, but are influenced by the glycan. If the procedure is to be used successfully for this purpose, peptide-specific antibodies should not have their binding diminished by deglycosylation of the antigen, and unaltered reactivity with antibodies has been reported for numerous glycoproteins after deglycosylation with TFMS.

Antibodies against CEA have been separated into those recognizing carbohydrate determinants and those recognizing the protein by this method. An antibody against the peptide of CEA, CP4, was unreactive with intact CEA prior to the removal of carbohydrate. When tested on human tissue, CP4 reacted only with CEA in colorectal tumours and did not react in immunohistochemistry with CEA in normal tissue or in stomach tumours, suggesting an alteration in the structure of the epitope in these tumours [86]. Antibodies against mucin proteins are also difficult to categorize, and progress has been made by deglycosylation of the mucins followed by antibody specificity studies [87–89]. This type of study has aided in the elucidation of multiple mucin genes that now account for the variants of the protein part of secreted and cell surface mucins [90,91]. Human bronchial mucins were almost completely deglycosylated with TFMS in 3 h at room temperature, and removal of fucose, galactose and *N*-acetylglucosamine was complete, whereas the serine/threonine-linked *N*-acetylgalactosamine was 75 % removed [40]. The mucin, MUC1, in secretions from COLO 205 cells was not recognized by the antibodies HMFG2 or SM3 without prior deglycosylation with TFMS [92]. These antibodies are raised against the tandem repeats of the polymorphic epithelial mucin, and the repeats are inaccessible in the heavily glycosylated mucin from COLO 205 cells, but are accessible to the antibodies and are recognized without prior deglycosylation in MUC1 from mammary carcinomas. Antibodies against sialyl Le^a recognize the carbohydrate of this mucin from COLO 205 cells. Another antibody against the protein, Ma552, reacted weakly with MUC1 from COLO 205 cells, but its binding was much increased after deglycosylation and it was shown to recognize a peptide determinant [93]. The efficacy of these antibodies in the treatment of colon or breast cancers has been shown to be influenced by the specificity of the antibodies for protein or carbohydrate [32,94]. Antibodies specific for the protein core should be more effective in treatment [32,91,94], as the MUC1 protein is less glycosylated in tumours and this permits an antibody of the correct specificity to discriminate between the mucin in tumours and in normal tissue.

Deglycosylation of mucins can also be an important tool in improving their immunogenicity if antibodies are not easily obtained against the intact mucin. Mucin glycans protect tumour mucins from the immune response and the weakly antigenic mucins often give rise to antibodies that recognize the carbohydrate. The antibodies will usually react with more than one mucin, as anti-carbohydrate antibodies tend to recognize the same determinant present on distinct core proteins. To overcome this issue, an antibody against tracheobronchial mucins was raised in rabbits by injection of TFMS-treated human mucin from patients with chronic bronchitis [95]. A deglycosylated mucin of 97 kDa was obtained from both the acidic and the neutral mucins and was enriched in serine, threonine and proline residues. The TFMS treatment removed all sialic acid, fucose, galactose and *N*-acetylglucosamine and all but traces of *N*-acetylgalactosamine. The antibodies against deglycosylated tracheobronchial mucins reacted strongly with the native mucins, as well as with mucins from human colon [95]. The antibody was used for screening of a cDNA library and resulted in the identification of a gene for MUC4 [96]. The antibody recognized the tandem repeats of MUC4 and could also be used to demonstrate the occurrence of MUC4 in pancreatic tumour cell lines [97,98]. Two mucins in these cell lines were recognized by the DU-PAN-2 antibody against the carbohydrate determinants of the mucins, but only one of these reacted with the tracheobronchial mucin antibody against the MUC4 core protein. Another mucin core protein antibody against pancreatic apomucin was generated using TFMS-deglycosylated mucin [99]. Gastric mucins were deglycosylated with TFMS to yield a core protein that was used as an immunogen for the generation of monoclonal antibodies that recognized core proteins in the small intestine and biliary tract [100]. This antibody was used to screen a cDNA expression library to yield a clone that contained the MUC5AC tandem repeat. The antibody recognized gastric epithelial cells and parietal cells, but not glandular cells, which is consistent with the distribution of MUC5AC.

Antibodies raised against zona pellucida glycoproteins after deglycosylation recognized the native zona pellucida glycoproteins and blocked binding of sperm [101]. The antibodies against porcine zona pellucida inhibited fertility in primates [101], and the deglycosylated proteins could be used to induce infertility actively by injection in rabbits [102].

The study of glycoprotein antigens after deglycosylation has also indicated the importance of some carbohydrate determinants in autoimmune diseases [11,103–106], and has revealed that carbohydrate determinants can be important in antigens used for diagnostic tests for circulating antibodies in autoimmune [107] and infectious diseases [22]. T-cell recognition of collagen in rheumatoid arthritis is dependent on carbohydrate epitopes, and removal of the hydroxylysine-linked sugars with TFMS indicated that the carbohydrate was necessary for reactivity of the T-cell clones [11,106]. Removal of keratan sulphate chains with TFMS increased reactivity of T-cells with aggrecan in mice [103] and in rheumatoid arthritis patients [104,105]. An antibody against the core protein of aggrecan was shown to be a useful marker for inflammatory disease, as synovial fluid from patients with degenerative joint diseases had increased proteoglycan fragments that could be detected with this antibody [107]. Tests of serum for antibodies against *T. gondii* used for diagnosis of toxoplasmosis can be performed using surface antigen I expressed by *P. pastoris*. Although the toxoplasma antigen lacks O-linked sugars, the recombinant molecule made for the test in yeast was glycosylated on hydroxyamino acids and the yeast oligosaccharides were recognized by some patient sera, complicating the diagnostic test. Deglycosylation of the antigen allowed the detection of

cross-reactivity of the patient serum with the toxoplasma protein [22].

Many allergens from insects and plants comprise glycoproteins that can stimulate the production of IgE. A molecular understanding of the allergenic epitopes is important for establishing the pathogenesis of the allergies. Allergens from a number of species contain fucose-linked α-1,3 and xylose-linked α-1,2 to *N*-acetylglucosamine that are unique to non-mammalian species, and are thought to be antigens for a common immune response to plant pollens, vegetable food and insect allergens. This was indicated for Japanese cedar (*Cryptomeria japonica*) [108], flour glycoproteins [109], the allergenic epitope from timothy grass (*Phleum pratense*) pollen [70] and olive tree (*Olea europaea*) pollen allergen [110], and is thought to be a key epitope from bee venom [16]. However, deglycosylation of bee venom with TFMS [16] indicated that the protein was more important for this response and that the deglycosylated antigen was capable of eliciting a full IgE response. In studies of plant allergens the reactivity between patient immune serum and protein or carbohydrate components of plant glycoproteins has been important for an understanding of the immune response. The timothy grass pollen antigen could be deglycosylated by TFMS, but not by N-glycanase, and yielded a 33 kDa protein that was tested for reactivity with antiserum from patients. Most sera bound to the protein core, rather than the oligosaccharides [70] which had previously been thought to be the antigenic determinant. In attempts to prepare specific antibodies against glycoproteins from plants, the immunogenicity of the carbohydrate structures that are foreign to mammals can give rise to anti-carbohydrate antibodies and these may not be useful, as they can react with multiple plant glycoproteins. By removing carbohydrates prior to immunization, a specific antibody against the protein core of β-1,3-glucanase was obtained and was useful in studies of enzyme function *in vivo* [111].

STRUCTURE–FUNCTION RELATIONSHIPS IN GLYCOPROTEINS

The effect of carbohydrate removal on the biological function of enzymes, hormones, lectins and enzyme inhibitors has been studied by several groups. In a detailed study of the deglycosylation of horseradish peroxidase, enzymic treatments with N-glycanase, endo F and endo H were unsuccessful in deglycosylating the enzyme [39]. The eight N-linked chains of the enzyme were effectively removed after TFMS treatment, yielding an enzyme of the correct deglycosylated molecular mass, and the enzyme was fully active after deglycosylation and purification. This result was in agreement with a study in which the protein produced in *Escherichia coli* without glycosylation was found to have about 50 % of the full activity of the intact enzyme [112]. The preferred scavenger for the deglycosylation reaction was toluene and the optimal conditions for obtaining the active enzyme were -10 °C for 5 min. Neutralization of the reaction was performed with Tris to avoid extraction of the haem group with pyridine. The occurrence of increased negative charge in the protein when longer reaction times were used was reported, but the source of the charge was not found and the deglycosylated protein that was purified and used for the enzyme assays did not contain the low pI material.

The glycoprotein hormones are not dependent on carbohydrates for receptor binding or immune reactivity, but the ability to stimulate a full response via the receptor is lost after deglycosylation [113,114]. Full receptor-binding activity has been demonstrated with several of the hormones using TFMS or HF for deglycosylation or using enzymes for removal of N-linked oligosaccharide chains [115]. In each case decreased signal transduction has been observed with the deglycosylated hormone. This is apparent in assays *in vitro* of cAMP production and steroidogenesis with the intact as compared with the deglycosylated hormones. These deglycosylated proteins are lutropin antagonists that bind receptor without giving rise to biological activity and can therefore inhibit the activity of lutropin [113].

Chicken ovomucoid, a trypsin inhibitor, was fully active after removal of sugars, but had increased susceptibility to proteolytic degradation and heat denaturation [116]. The deglycosylated inhibitor had unchanged activity in inhibiting trypsin and was recognized by antibodies against the native molecule. The sugars of ovomucoid allow formation of a stable 1:1 complex with trypsin and thus are a critical part of the molecule for enzyme inhibition. The glycoprotein without carbohydrate was degraded by excess trypsin [116] unlike the native molecule, indicating that glycosylation of this inhibitor is imperative for protecting the protein inhibitor from its target protease. Dabich et al. [117] found that turkey ovomucoid was rendered less effective as a trypsin inhibitor, an activity dependent on the second domain of the molecule, but had increased activity towards chymotrypsin, an activity that resides in the third domain, but may be constrained by glycosylation of the second domain. The cysteine protease inhibitor thiostatin depended on glycosylation for its inhibitory activity [118], but retained its reactivity with antibodies to the native protein.

The interaction of mesangial cells with fibronectin is inhibited by the major heparan sulphate proteoglycan of the glomerular basement membrane [61]. This proteoglycan, perlecan, reduced the avidity of the $\alpha 5\beta 1$ integrin for fibronectin in mesangial cell focal adhesions. Complete removal of the carbohydrate from perlecan did not decrease inhibition and the authors concluded that the core protein, and not the heparan sulphate chains, was required to inhibit a high-affinity interaction of fibronectin with its integrin receptor [61]. However, Lewis lung carcinoma cells contained a proteoglycan that interacted with fibronectin and depended upon the heparan sulphate chains for the high-affinity interaction, as shown by deglycosylation of the proteoglycan with TFMS and with glycosaminoglycan-degrading enzymes [36]. The proteoglycan was inserted into the membrane of the cells and the affinity of its heparan sulphate chains for fibronectin accounted for binding of the cells to the extracellular matrix.

Cobra venom factor retained full biological activity upon complete removal of carbohydrate by TFMS [119]. The peptide of potato lectin was recovered intact after deglycosylation with TFMS and was used to determine the function of the carbohydrate in the lectin's activity. The peptide retained full lectin activity (binding to chitobiose) and haemagglutinating activity [18]. The molecular mass of the deglycosylated lectin measured by MALDI–TOF MS was 31 253 Da (native lectin = 65 492 Da) [120]. The potato lectin shows similarity to the snake venom disintegrins, indicating that these proteins do not need carbohydrate to recognize their ligands or to carry out their functions.

SUMMARY

Deglycosylation of glycoproteins with TFMS is a powerful technique that can be used to determine the size of the protein portion of a glycoprotein and to reduce all glycosylation variants to a single parent molecule. The protein size correlates with the predicted size of known genes, and thus deglycosylation with TFMS has made it possible to correlate numerous glycoforms of a glycoprotein to a common protein core. Sequencing of proteins after chemical deglycosylation has been performed in numerous studies in which analysis of *N*-acetylglucosamine–asparagine

for N-linked oligosaccharides and *N*-acetylgalactosamine–serine/threonine for O-linked oligosaccharides has made possible the identification of glycosylation sites. A scavenger should be included in the reaction to avoid alterations in individual amino acids. A decrease in molecular mass by TFMS treatment provides a specific means of demonstrating the occurrence of glycans, because other protein modifications are not affected. The method has allowed the elucidation of glycoforms in the processing of newly synthesized glycoproteins. The deglycosylation of glycoproteins in functional studies has provided valuable information on the role of sugars in the biological activity of hormones, lectins, enzymes, protease inhibitors and extracellular matrix components. These studies must be interpreted with caution, but retention of activity after deglycosylation has indicated that the biological and immunochemical activity of numerous glycoproteins resides in the protein part of the molecule. The removal of carbohydrate from glycoproteins by TFMS has permitted these advances in glycobiology by providing a means of dissecting glycan from protein, without degradation of the protein.

I thank A. J. Kassenaar for the artwork in the figures.

REFERENCES

1 Edge, A. S. B., Faltynek, C. R., Hof, L., Reichert, Jr, L. E. and Weber, P. (1981) Deglycosylation of glycoproteins by trifluoromethanesulfonic acid. Anal. Biochem. **118**, 131–137

2 Florman, H. M. and Wassarman, P. M. (1985) O-linked oligosaccharides of mouse egg ZP3 account for its sperm receptor activity. Cell (Cambridge, Mass.) **41**, 313–324

3 Herzberg, V. L., Grigorescu, F., Edge, A. S. B., Spiro, R. G. and Kahn, C. R. (1985) Characterization of insulin receptor carbohydrate by comparison of chemical and enzymatic deglycosylation. Biochem. Biophys. Res. Commun. **129**, 789–796

4 Karp, D. R., Parker, K. L., Shreffler, D. C., Slaughter, C. and Capra, J. D. (1982) Amino acid sequence homologies and glycosylation differences between the fourth component of murine complement and sex-limited protein. Proc. Natl. Acad. Sci. U.S.A. **79**, 6347–6349

5 Harrison, R. A. (1983) Human C1 inhibitor: improved isolation and preliminary structural characterization. Biochemistry **22**, 5001–5007

6 Paxton, R. J., Mooser, G., Pande, H., Lee, T. D. and Shively, J. E. (1987) Sequence analysis of carcinoembryonic antigen: identification of glycosylation sites and homology with the immunoglobulin supergene family. Proc. Natl. Acad. Sci. U.S.A. **84**, 920–924

7 Naim, H. Y., Sterchi, E. E. and Lentze, M. J. (1988) Biosynthesis of the human sucrase-isomaltase complex. Differential O-glycosylation of the sucrase subunit correlates with its position within the enzyme complex. J. Biol. Chem. **263**, 7242–7253

8 Edge, A. S. B. and Spiro, R. G. (1987) Selective deglycosylation of the heparan sulfate proteoglycan of bovine glomerular basement membrane and identification of the core protein. J. Biol. Chem. **262**, 6893–6898

9 Jost, C. J., Funderburgh, J. L., Mann, M., Hassell, J. R. and Conrad, G. W. (1991) Cell-free translation and characterization of corneal keratan sulfate proteoglycan core proteins. J. Biol. Chem. **266**, 13336–13341

10 Lories, V., Cassiman, J. J., Van den Berghe, H. and David, G. (1989) Multiple distinct membrane heparan sulfate proteoglycans in human lung fibroblasts. J. Biol. Chem. **264**, 7009–7016

11 Michaelsson, E., Malmstrom. V., Reis. S., Engstrom, A., Burkhardt, H. and Holmdahl, R. (1994) T cell recognition of carbohydrates on type II collagen. J. Exp. Med. **180**, 745–749

12 Marton, L. S. and Stefansson, K. (1986) Exposure of binding sites for antibodies and concanavalin A on collagen by solubilization in hot urea. An immunoblot analysis. J. Immunol. Methods **91**, 187–194

13 Sylvestre, P., Couture-Tosi, E. and Mock, M. (2002) A collagen-like surface glycoprotein is a structural component of the *Bacillus anthracis* exosporium. Mol. Microbiol. **45**, 169–178

14 Dissous, C., Grzych, J. M. and Capron, A. (1986) *Schistosoma mansoni* shares a protective oligosaccharide epitope with freshwater and marine snails. Nature (London) **323**, 443–445

15 Kramerova, I. A., Kawaguchi, N., Fessler, L. I., Nelson, R. E., Chen, Y., Kramerov, A. A., Kusche-Gullberg, M., Kramer, J. M., Ackley, B. D., Sieron, A. L. et al. (2000) Papilin in development: a pericellular protein with a homology to the ADAMTS metalloproteinases. Development **127**, 5475–5485

16 Okano, M., Nishizaki. K., Satoskar, A. R., Yoshino, T., Masuda, Y. and Harn, Jr, D. A. (1999) Involvement of carbohydrate on phospholipase A2, a bee-venom allergen, in *in vivo* antigen-specific IgE synthesis in mice. Allergy **54**, 811–818

17 Jarchow, J., Fritz, J., Anselmetti, D., Calabro, A., Hascall, V. C., Gerosa, D., Burger, M. M. and Fernandez-Busquets, X. (2000) Supramolecular structure of a new family of circular proteoglycans mediating cell adhesion in sponges. J. Struct. Biol. **132**, 95–105

18 Desai, N. N., Allen, A. K. and Neuberger, A. (1983) The properties of potato (*Solanum tuberosum*) lectin after deglycosylation by trifluoromethanesulphonic acid. Biochem. J. **211**, 273–276

19 Lind, J. L., Bacic, A., Clarke, A. E. and Anderson, M. A. (1994) A style-specific hydroxyproline-rich glycoprotein with properties of both extensins and arabinogalactan proteins. Plant J. **6**, 491–502

20 Bosch, M., Knudsen, J. S., Derksen, J. and Mariani, C. (2001) Class III pistil-specific extensin-like proteins from tobacco have characteristics of arabinogalactan proteins. Plant Physiol. **125**, 2180–2188

21 Takeichi, T., Takeuchi, J., Kaneko, T. and Kawasaki, S. (1998) Purification and characterization of a galactose-rich basic glycoprotein in tobacco. Plant Physiol. **116**, 477–483

22 Letourneur, O., Gervasi, G., Gaia, S., Pages, J., Watelet, B. and Jolivet, M. (2001) Characterization of *Toxoplasma gondii* surface antigen 1 (SAG1) secreted from *Pichia pastoris*: evidence of hyper O-glycosylation. Biotechnol. Appl. Biochem. **33**, 35–45

23 Christian, R., Schulz, G., Schuster-Kolbe, J., Allmaier, G., Schmid, E. R., Sleytr, U. B. and Messner, P. (1993) Complete structure of the tyrosine-linked saccharide moiety from the surface layer glycoprotein of *Clostridium thermohydrosulfuricum* S102-70. J. Bacteriol. **175**, 1250–1256

24 Schaffer, C., Wugeditsch, T., Kahlig, H., Scheberl, A., Zayni, S. and Messner, P. (2002) The surface layer (S-layer) glycoprotein of *Geobacillus stearothermophilus* NRS 2004/3a. Analysis of its glycosylation. J. Biol. Chem. **277**, 6230–6239

25 Lower, B. H. and Kennelly, P. J. (2002) The membrane-associated protein-serine/threonine kinase from *Sulfolobus solfataricus* is a glycoprotein. J. Bacteriol. **184**, 2614–2619

26 Jubilut, G. N., Cilli, E. M., Tominaga, M., Miranda, A., Okada, Y. and Nakaie, C. R. (2001) Evaluation of the trifluoromethanosulfonic acid/trifluoroacetic acid/thioanisole cleavage procedure for application in solid-phase peptide synthesis. Chem. Pharm. Bull. (Tokyo) **49**, 1089–1092

27 Thompson, P. E., Keah, H. H., Gomme, P. T., Stanton, P. G. and Hearn, M. T. (1995) Synthesis of peptide amides using Fmoc-based solid-phase procedures on 4-methylbenzhydrylamine resins. Int. J. Pept. Protein Res. **46**, 174–180

28 Lebl, M., Pires, J., Poncar, P. and Pokorny, V. (1999) Evaluation of gaseous hydrogen fluoride as a convenient reagent for parallel cleavage from the solid support. J. Comb. Chem. **1**, 474–479

29 Torok, B., Busci, I., Surya Prakash, G. K. and Olah, G. A. (2002) Deprotection and cleavage of peptide bound to Merrifield resin by stable dimethyl ether-poly(hydrogen fluoride) (DMEPHF) complex. A new and convenient reagent for peptide chemistry. Chem. Commun. **23**, 2882–2883

30 Woodward, H. D., Ringler, N. J., Selvakumar, R., Simet, I. M., Bhavanandan, V. P. and Davidson, E. A. (1987) Deglycosylation studies on tracheal mucin glycoproteins. Biochemistry **26**, 5315–5322

31 Muller, S., Goletz, S., Packer, N., Gooley, A., Lawson, A. M. and Hanisch, F. G. (1997) Localization of O-glycosylation sites on glycopeptide fragments from lactation-associated MUC1. All putative sites within the tandem repeat are glycosylation targets *in vivo*. J. Biol. Chem. **272**, 24780–24793

32 Lloyd, K. O., Burchell, J., Kudryashov, V., Yin, B. W. and Taylor-Papadimitriou, J. (1996) Comparison of O-linked carbohydrate chains in MUC-1 mucin from normal breast epithelial cell lines and breast carcinoma cell lines. Demonstration of simpler and fewer glycan chains in tumor cells. J. Biol. Chem. **271**, 33325–33334

33 Hanisch, F. G., Uhlenbruck, G., Peter-Katalinic, J., Egge, H., Dabrowski, J. and Dabrowski, U. (1989) Structures of neutral O-linked polylactosaminoglycans on human skim milk mucins. A novel type of linearly extended poly-*N*-acetyl-lactosamine backbones with Gal β(1-4)GlcNAc β(1-6) repeating units. J. Biol. Chem. **264**, 872–883

34 Gerken, T. A., Gupta, R. and Jentoft, N. (1992) A novel approach for chemically deglycosylating O-linked glycoproteins. The deglycosylation of submaxillary and respiratory mucins. Biochemistry **31**, 639–648

35 Hagen, S. G., Michael, A. F. and Butkowski, R. J. (1993) Immunochemical and biochemical evidence for distinct basement membrane heparan sulfate proteoglycans. J. Biol. Chem. **268**, 7261–7269

36 Itano, N., Oguri, K., Nakanishi, H. and Okayama, M. (1993) Membrane-intercalated proteoglycan of a stroma-inducing clone from Lewis lung carcinoma binds to fibronectin via its heparan sulfate chains. J. Biochem. (Tokyo) **114**, 862–873

37 Engelmann, S. and Schwartz-Albiez, R. (1997) Differential release of proteoglycans during human B lymphocyte maturation. Carbohydr. Res. **302**, 85–95

38 Forbes, J. and Beeley, J. G. (1989) Deglycosylation of glycoproteins with trifluoromethanesulphonic acid – use of water-soluble scavengers. Biochem. Soc. Trans. **17**, 737

39 Tams, J. W. and Welinder, K. G. (1995) Mild chemical deglycosylation of horseradish peroxidase yields a fully active, homogeneous enzyme. Anal. Biochem. **228**, 48–55

40 Marianne, T., Perini, J. M., Houvenaghel, M. C., Tramu, G., Lamblin, G. and Roussel, P. (1986) Action of trifluoromethanesulfonic acid on highly glycosylated regions of human bronchial mucins. Carbohydr. Res. **151**, 7–19

41 Maley, F., Trimble, R. B., Tarentino, A. L. and Plummer, Jr, T. H. (1989) Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. Anal. Biochem. **180**, 195–204

42 Dwek, R. A., Edge, C. J., Harvey, D. J., Wormald, M. R. and Parekh, R. B. (1993) Analysis of glycoprotein-associated oligosaccharides. Annu. Rev. Biochem. **62**, 65–100

43 Bhattacharyya, S. N., Enriquez, Sr, J. I. and Manna, B. (1990) Deglycosylation of neutral and acidic human colonic mucin. Inflammation **14**, 93–107

44 Bell, S. L., Khatri, I. A., Xu, G. and Forstner, J. F. (1998) Evidence that a peptide corresponding to the rat Muc2 C-terminus undergoes disulphide-mediated dimerization. Eur. J. Biochem. **253**, 123–131

45 Gerken, T. A., Owens, C. L. and Pasumarthy, M. (1997) Determination of the site-specific O-glycosylation pattern of the porcine submaxillary mucin tandem repeat glycopeptide. Model proposed for the polypeptide:GalNAc transferase peptide binding site. J. Biol. Chem. **272**, 9709–9719

46 Hascall, V. C., Calabro, A., Midura, R. J. and Yanagashita, M. (1994) Isolation and characterization of proteoglycans. Methods Enzymol. **230**, 390–417

47 Haltiwanger, R. S. (2002) Regulation of signal transduction pathways in development by glycosylation. Curr. Opin. Struct. Biol. **12**, 593–598

48 Desarnaud, F., Marie, J., Lombard, C., Larguier, R., Seyer, R., Lorca, T., Jard, S. and Bonnafous, J. C. (1993) Deglycosylation and fragmentation of purified rat liver angiotensin II receptor: application to the mapping of hormone-binding domains. Biochem. J. **289**, 289–297

49 Funderburgh, J. L. and Conrad, G. W. (1990) Isoforms of corneal keratan sulfate proteoglycan. J. Biol. Chem. **265**, 8297–8303

50 Romain, F., Horn, C., Pescher, P., Namane, A., Riviere, M., Puzo, G., Barzu, O. and Marchal, G. (1999) Deglycosylation of the 45/47-kilodalton antigen complex of *Mycobacterium tuberculosis* decreases its capacity to elicit *in vivo* or *in vitro* cellular immune responses. Infect. Immun. **67**, 5567–5572

51 Litscher, E. S., Qi, H. and Wassarman, P. M. (1999) Mouse zona pellucida glycoproteins mZP2 and mZP3 undergo carboxy-terminal proteolytic processing in growing oocytes. Biochemistry **38**, 12280–12287

52 de Vocht, M. L., Scholtmeijer, K., van der Vegte, E. W., de Vries, O. M., Sonveaux, N., Wosten, H. A., Ruysschaert, J. M., Hadziloannou, G., Wessels, J. G. and Robillard, G. T. (1998) Structural characterization of the hydrophobin SC3, as a monomer and after self-assembly at hydrophobic/hydrophilic interfaces. Biophys. J. **74**, 2059–2068

53 Hefta, S. A., Paxton, R. J. and Shively, J. E. (1990) Sequence and glycosylation site identity of two distinct glycoforms of nonspecific cross-reacting antigen as demonstrated by sequence analysis and fast atom bombardment mass spectrometry. J. Biol. Chem. **265**, 8618–8626

54 Enghild, J. J., Thogersen, I. B., Pizzo, S. V. and Salvesen, G. (1989) Analysis of inter-α-trypsin inhibitor and a novel trypsin inhibitor, pre-α-trypsin inhibitor, from human plasma. Polypeptide chain stoichiometry and assembly by glycan. J. Biol. Chem. **264**, 15975–15981

55 Hedrick, J. L. and Wardrip, N. J. (1987) On the macromolecular composition of the zona pellucida from porcine oocytes. Dev. Biol. **121**, 478–488

56 Malette, B. and Bleau, G. (1993) Biochemical characterization of hamster oviductin as a sulphated zona pellucida-binding glycoprotein. Biochem. J. **295**, 437–445

57 Edge, A. S., Kahn, C. R. and Spiro, R. G. (1990) Insulin receptor carbohydrate units contain poly-*N*-acetyllactosamine chains. Endocrinology **127**, 1887–1895

58 van den Heuvel, L. P., van den Born, J., Veerkamp, J. H., van de Velden, T. J., Schenkels, L., Monnens, L. A., Schroder, C. H. and Berden, J. H. (1990) Heparan sulfate proteoglycan from human tubular basement membrane. Comparison with this component from the glomerular basement membrane. Biochim. Biophys. Acta **1025**, 67–76

59 Groffen, A. J., Ruegg, M. A., Dijkman, H., van de Velden, T. J., Buskens, C. A., van den Born, J., Assmann, K. J., Monnens, L. A., Veerkamp, J. H. and van den Heuvel, L. P. (1998) Agrin is a major heparan sulfate proteoglycan in the human glomerular basement membrane. J. Histochem. Cytochem. **46**, 19–27

60 Yard, B. A., Kahlert, S., Engelleiter, R., Resch, S., Waldherr, R., Groffen, A. J., van den Heuvel, L. P., van der Born, J, Berden, J. H., Kroger, S. et al. (2001) Decreased glomerular expression of agrin in diabetic nephropathy and podocytes, cultured in high glucose medium. Exp. Nephrol. **9**, 214–222

61 Gauer, S., Schulze-Lohoff, E., Schleicher, E. and Sterzel, R. B. (1996) Glomerular basement membrane-derived perlecan inhibits mesangial cell adhesion to fibronectin. Eur. J. Cell Biol. **70**, 233–242

62 Soroka, C. J. and Farquhar, M. G. (1991) Characterization of a novel heparan sulfate proteoglycan found in the extracellular matrix of liver sinusoids and basement membranes. J. Cell Biol. **113**, 1231–1241

63 Naim, H. Y. and Lentze, M. J. (1992) Impact of O-glycosylation on the function of human intestinal lactase-phlorizin hydrolase. Characterization of glycoforms varying in enzyme activity and localization of *O*-glycoside addition. J. Biol. Chem. **267**, 25494–25504

64 Naim, H. Y., Joberty, G., Alfalah, M. and Jacob, R. (1999) Temporal association of the N- and O-linked glycosylation events and their implication in the polarized sorting of intestinal brush border sucrase-isomaltase, aminopeptidase N, and dipeptidyl peptidase IV. J. Biol. Chem. **274**, 17961–17967

65 Tague, B. W. and Chrispeels, M. J. (1987) The plant vacuolar protein, phytohemagglutinin, is transported to the vacuole of transgenic yeast. J. Cell Biol. **105**, 1971–1979

66 Sturm, A., Voelker, T. A., Herman, E. M. and Chrispeels, M. J. (1988) Correct glycosylation, Golgi-processing, and targeting to protein bodies of the vacuolar protein phytohemagglutinin in transgenic tobacco. Planta **175**, 170–183

67 Sonnewald, U., Sturm, A., Chrispeels, M. J. and Willmitzer, L. (1989) Targeting and glycosylation of patatin the major potato-tuber protein in leaves of transgenic tobacco. Planta **179**, 171–180

68 Hunt, D. C. and Chrispeels, M. J. (1991) The signal peptide of a vacuolar protein is necessary and sufficient for the efficient secretion of a cytosolic protein. Plant Physiol. **96**, 18–25

69 Schnell, D. J., Hori, K., Herrmann, S. M., Gegg, C. V. and Etzler, M. E. (1994) Biosynthesis of DB58 lectin in cell suspension cultures of *Dolichos biflorus*. Arch. Biochem. Biophys. **310**, 229–235

70 Petersen, A., Becker, W. M., Moll, H., Blumke, M. and Schlaak, M. (1995) Studies on the carbohydrate moieties of the timothy grass pollen allergen Phl pI. Electrophoresis **16**, 869–875

71 Altman, E., Schaffer, C., Brisson, J. R. and Messner, P. (1995) Characterization of the glycan structure of a major glycopeptide from the surface layer glycoprotein of *Clostridium thermosaccharolyticum* E207-71. Eur. J. Biochem. **229**, 308–315

72 Dhiman, N., Verma, I. and Khuller, G. K. (1997) Cellular immune responses to cell wall peptidoglycan associated protein antigens in tuberculosis patients and healthy subjects. Microbiol. Immunol. **41**, 495–502

73 Dhiman, N. and Khuller, G. K. (1997) Immunoprophylactic properties of 71-kDa cell wall-associated protein antigen of *Mycobacterium tuberculosis* H37Ra. Med. Microbiol. Immunol. (Berlin) **186**, 45–51

74 Hirschfield, G. R., McNeil, M. and Brennan, P. J. (1990) Peptidoglycan-associated polypeptides of *Mycobacterium tuberculosis*. J. Bacteriol. **172**, 1005–1013

75 Aono, R., Ito, M. and Horikoshi, K. (1993) Occurrence of teichuronopeptide in the cell walls of group 2 alkaliphilic *Bacillus* spp. J. Gen. Microbiol. **139**, 2739–2744

76 Knirel, Y. A. and Pereplov, A. V. (2002) Trifluoromethanesulfonic acid: a useful reagent for the solvolytic cleavage of glycosidic linkages in structural analysis of bacterial polysaccharides. Aust. J. Chem. **55**, 69–72

77 Perepelov, A. V., Babicka, D., Senchenkova, S. N., Shashkov, A. S., Moll, H., Rozalski, A., Zahringer, U. and Knirel Y. A. (2001) Structure of the O-specific polysaccharide of *Proteus vulgaris* O4 containing a new component of bacterial polysaccharides, 4,6-dideoxy-4-{*N*-[(*R*)-3-hydroxybutyryl]-L-alanyl}amino-D-glucose. Carbohydr. Res. **331**, 195–202

78 Kocharova, N. A., Knirel, Y. A., Jansson, P. and Weintraub, A. (2001) Structure of the O-specific polysaccharide of *Vibrio cholerae* O9 containing 2-acetamido-2-deoxy-D-galacturonic acid. Carbohydr. Res. **332**, 279–284

79 Horvath, E., Edwards, A. M., Bell, J. C. and Braun, P. E. (1989) Chemical deglycosylation on a micro-scale of membrane glycoproteins with retention of phosphoryl-protein linkages. J. Neurosci. Res. **24**, 398–401

80 Yun, H. Y., Keutmann, H. T. and Eipper, B. A. (1994) Alternative splicing governs sulfation of tyrosine or oligosaccharide on peptidylglycine α-amidating monooxygenase. J. Biol. Chem. **269**, 10946–10955

81 Kohli, R., Chadha, N. and Muralidhar, K. (1988) Presence of tyrosine-*O*-sulfate in sheep pituitary prolactin. FEBS Lett. **242**, 139–143

82 Haraguchi, T., Fisher, S., Olofsson, S., Endo, T., Groth, D., Tarentino, A., Borchelt, D. R., Teplow, D., Hood, L., Burlingame, A., et al. (1989) Asparagine-linked glycosylation of the scrapie and cellular prion proteins. Arch. Biochem. Biophys. **274**, 1–13

83 Deddish, P. A., Skidgel, R. A., Kriho, V. B., Li, X. Y., Becker, R. P. and Erdos, E. G. (1990) Carboxypeptidase M in Madin–Darby canine kidney cells. Evidence that carboxypeptidase M has a phosphatidylinositol glycan anchor. J. Biol. Chem. **265**, 15083–15089

84 Hortsch, M. and Goodman, C. S. (1990) *Drosophila* fasciclin I, a neural cell adhesion molecule, has a phosphatidylinositol lipid membrane anchor that is developmentally regulated. J. Biol. Chem. **265**, 15104–15109

85 Sampath, T. K., Coughlin, J. E., Whetstone, R. M., Banach, D., Corbett, C., Ridge, R. J., Ozkaynak, E., Oppermann, H. and Rueger, D. C. (1990) Bovine osteogenic protein is composed of dimers of OP-1 and BMP-2A, two members of the transforming growth factor-β superfamily. J. Biol. Chem. **265**, 13198–13205

86 Horan Hand, P., Simpson, J. F., Kuroki, M., Parker, R. and Schlom, J. (1992) Reactivities of an anti-CEA peptide monoclonal antibody. Int. J. Biol. Markers **7**, 1–15

87 Stahli, C., Caravatti, M., Aeschbacher, M., Kocyba, C., Takacs, B. and Carmann, H. (1988) Mucin-like carcinoma-associated antigen defined by three monoclonal antibodies against different epitopes. Cancer Res. **48**, 6799–6802

88 Sheer, D. G., Schlom, J. and Cooper, H. L. (1988) Purification and composition of the human tumor-associated glycoprotein (TAG-72) defined by monoclonal antibodies CC49 and B72.3. Cancer Res. **48**, 6811–6818

89 Yin, B. W. and Lloyd, K. O. (2001) Molecular cloning of the CA125 ovarian cancer antigen: identification as a new mucin, MUC16. J. Biol. Chem. **276**, 27371–27375

90 Thornton, D. J., Howard, M., Devine, P. L. and Sheehan, J. K. (1995) Methods for separation and deglycosylation of mucin subunits. Anal. Biochem. **227**, 162–167

91 Dalziel, M., Whitehouse, C., McFarlane, I., Brockhausen, I., Gschmeissner, S., Schwientek, T., Clausen, H., Burchell, J. M. and Taylor-Papadimitriou, J. (2001) The relative activities of the C2GnT1 and ST3Gal-I glycosyltransferases determine O-glycan structure and expression of a tumor-associated epitope on MUC1. J. Biol. Chem. **276**, 11007–11015

92 Baeckstrom, D., Hansson, G. C., Nilsson, O., Johansson, C., Gendler, S. J. and Lindholm, L. (1991) Purification and characterization of a membrane-bound and a secreted mucin-type glycoprotein carrying the carcinoma-associated sialyl-Le^a epitope on distinct core proteins. J. Biol. Chem. **266**, 21537–21547

93 Baeckstrom, D., Nilsson, O., Price, M. R., Lindholm, L. and Hansson, G. C. (1993) Discrimination of MUC1 mucins from other sialyl-Le^a-carrying glycoproteins produced by colon carcinoma cells using a novel monoclonal antibody. Cancer Res. **53**, 755–761

94 Reis, C. A., David, L., Seixas, M, Burchell, J. and Sobrinho-Simoes, M. (1998) Expression of fully and under-glycosylated forms of MUC1 mucin in gastric carcinoma. Int. J. Cancer **79**, 402–410

95 Bhattacharyya, S. N., Veit, B. C., Manna, B., Enriquez, J. I., Walker, M. P., Khorrami, A. M. and Kaufman, B. (1990) Neutral and acidic human tracheobronchial mucin. Isolation and characterization of core protein. Inflammation **14**, 355–373

96 Choudhury, A., Moniaux, N., Winpenny, J. P., Hollingsworth, M. A., Aubert, J. P. and Batra, S. K. (2000) Human MUC4 mucin cDNA and its variants in pancreatic carcinoma. J. Biochem. (Tokyo) **128**, 233–243

97 Hollingsworth, M. A., Strawhecker, J. M., Caffrey, T. C. and Mack, D. R. (1994) Expression of MUC1, MUC2, MUC3 and MUC4 mucin mRNAs in human pancreatic and intestinal tumor cell lines. Int. J. Cancer **57**, 198–203

98 Khorrami, A. M., Choudhury, A., Andrianifahanana, M., Varshney, G. C., Bhattacharyya, S. N., Hollingsworth, M. A., Kaufman, B. and Batra, S. K. (2002) Purification and characterization of a human pancreatic adenocarcinoma mucin. J. Biochem. (Tokyo) **131**, 21–29

99 Lan, M. S., Batra, S. K., Qi, W. N., Metzgar, R. S. and Hollingsworth, M. A. (1990) Cloning and sequencing of a human pancreatic tumor mucin cDNA. J. Biol. Chem. **265**, 15294–15299

100 Sotozono, M., Okada, Y., Sasagawa, T., Nakatou, T., Yoshida, A., Yokoi, T., Kubota, M. and Tsuji, T. (1996) Novel monoclonal antibody, SO-MU1, against human gastric MUC5AC apomucin. J. Immunol. Methods **192**, 87–96

101 Henderson, C. J., Hulme, M. J. and Aitken, R. J. (1987) Analysis of the biological properties of antibodies raised against intact and deglycosylated porcine zonae pellucidae. Gamete Res. **16**, 323–341

102 Bhatnagar, P., Gupta, S. K. and Sehgal, S. (1992) Immunogenicity of deglycosylated zona pellucida antigens and their inhibitory effects on fertility in rabbits. Int. J. Fertil. **37**, 53–63

103 Goodacre, J. A., Middleton, S., Lynn, S., Ross, D. A. and Pearson, J. (1993) Human cartilage aggrecan CS1 region contains cryptic T-cell recognition sites. Immunology **78**, 586–591

104 Goodstone, N. J., Doran, M. C., Hobbs, R. N., Butler, R. C., Dixey, J. J. and Ashton, B. A. (1996) Cellular immunity to cartilage aggrecan core protein in patients with rheumatoid arthritis and non-arthritic controls. Ann. Rheum. Dis. **55**, 40–46

105 Guerassimov, A., Zhang, Y., Banerjee, S., Cartman, A., Leroux, J. Y., Rosenberg, L. C., Esdaile, J., Fitzcharles, M. A. and Poole, A. R. (1998) Cellular immunity to the G1 domain of cartilage proteoglycan aggrecan is enhanced in patients with rheumatoid arthritis but only after removal of keratan sulfate. Arthritis Rheum. **41**, 1019–1025

106 Backlund, J., Carlsen, S., Hoger, T., Holm, B., Fugger, L., Kihlberg, J., Burkhardt, H. and Holmdahl, R. (2002) Predominant selection of T cells specific for the glycosylated collagen type II epitope (263–270) in humanized transgenic mice and in rheumatoid arthritis. Proc. Natl. Acad. Sci. U.S.A. **99**, 9960–9965

107 Fischer, D. C., Kolbe-Busch, S., Stocker, G., Hoffmann, A. and Haubeck, H. D. (1994) Development of enzyme immunoassays specific for keratan sulphate- and core-protein-epitopes of the large aggregating proteoglycan from human articular cartilage. Eur. J. Clin. Chem. Clin. Biochem. **32**, 285–291

108 Ogawa, H., Hijikata, A., Amano, M., Kojima, K., Fukushima, H., Ishizuka, I., Kurihara, Y. and Matsumoto, I. (1996) Structures and contribution to the antigenicity of oligosaccharides of Japanese cedar (*Cryptomeria japonica*) pollen allergen Cry j I: relationship between the structures and antigenic epitopes of plant N-linked complex-type glycans. Glycoconj. J. **13**, 555–566

109 Garcia-Casado, G., Sanchez-Monge, R., Chrispeels, M. J., Armentia, A., Salcedo, G. and Gomez, L. (1996) Role of complex asparagine-linked glycans in the allergenicity of plant glycoproteins. Glycobiology **6**, 471–477

110 Batanero, E., Villalba, M. and Rodriguez, R. (1994) Glycosylation site of the major allergen from olive tree pollen. Allergenic implications of the carbohydrate moiety. Mol. Immunol. **31**, 31–37

111 Kurosaki, F. (1997) Preparation of monoclonal antibody against plant extracellular glycoprotein with deglycosylated peptide as antigen. Anal. Biochem. **244**, 170–172

112 Smith, A. T., Santama, N., Dacey, S., Edwards, M., Bray, R. C., Thorneley, R. N. and Burke, J. F. (1990) Expression of a synthetic gene for horseradish peroxidase C in *Escherichia coli* and folding and activation of the recombinant enzyme with Ca^{2+} and heme. J. Biol. Chem. **265**, 13335–13343

113 Sairam, M. R. and Bhargavi, G. N. (1985) A role for glycosylation of the α subunit in transduction of biological signal in glycoprotein hormones. Science (Washington, D.C.) **229**, 65–67

114 Sojar, H. T. and Bahl, O. P. (1987) A chemical method for the deglycosylation of proteins. Arch. Biochem. Biophys. **259**, 52–57

115 Thotakura, N. R., LiCalzi, L. and Weintraub, B. D. (1990) The role of carbohydrate in thyrotropin action assessed by a novel approach using enzymatic deglycosylation. J. Biol. Chem. **265**, 11527–11534

116 Gu, J. X., Matsuda, T., Nakamura, R., Ishiguro, H., Ohkubo, I., Sasaki, M. and Takahashi, N. (1989) Chemical deglycosylation of hen ovomucoid: protective effect of carbohydrate moiety on tryptic hydrolysis and heat denaturation. J. Biochem. (Tokyo) **106**, 66–70

117 Dabich, D., Yurewicz, E. C. and Battel, V. A. (1993) Deglycosylation with trifluoromethanesulfonic acid differentially affects inhibitor activities of turkey ovomucoid. Biochim. Biophys. Acta **1164**, 47–53

118 Rusiniak, M. E., Bedi, G. S. and Back, N. (1991) Role of carbohydrate in rat plasma thiostatin: deglycosylation destroys cysteine proteinase inhibition activity. Biochem. Biophys. Res Commun. **179**, 927–932

119 Gowda, D. C., Petrella, E. C., Raj, T. T., Bredehorst, R. and Vogel, C. W. (1994) Immunoreactivity and function of oligosaccharides in cobra venom factor. J. Immunol. **152**, 2977–2986

120 Allen, A. K., Bolwell, G. P., Brown, D. S., Sidebottom, C. and Slabas, A. R. (1996) Potato lectin: a three-domain glycoprotein with novel hydroxyproline-containing sequences and sequence similarities to wheat-germ agglutinin. Int. J. Biochem. Cell Biol. **28**, 1285–1291

121 Edge, A. S. B. and Spiro, R. G. (1990) Characterization of novel sequences containing 3-O-sulfated glucosamine in glomerular basement membrane heparan sulfate and localization of sulfated disaccharides to a peripheral domain. J. Biol. Chem. **265**, 15874–15881

Received 8 May 2003/29 July 2003; accepted 15 September 2003
Published as BJ Immediate Publication 15 September 2003, DOI 10.1042/BJ20030673